John von Neumann Institute for Computing

NewsGrid: Infrastructure and Services

S. Geisler, G.R. Joubert

published in

Parallel Computing:
Current & Future Issues of High-End Computing,
Proceedings of the International Conference ParCo 2005,
G.R. Joubert, W.E. Nagel, F.J. Peters, O. Plata, P. Tirado, E. Zapata (Editors),
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **33**, ISBN 3-00-017352-8, pp. 105-112, 2006.

http://www.fz-juelich.de/nic-series/volume33

NewsGrid: Infrastructure and Services

S. Geisler[a], G.R. Joubert[a]

[a]Department of Computer Science, Clausthal University of Technology, Julius-Albert-Straße 4, 38678 Clausthal-Zellerfeld, Germany, E-mail: {geisler, joubert}@informatik.tu-clausthal.de

1. Introduction

TV stations need access to news video material in order to compile news casts. Such material is collected twenty-four hours per day seven days per week from different sources such as news agencies and own correspondents in analogue and/or digital form via various networks (telephone, satellite transmissions, mail, etc.). It cannot be predicted in advance which of this material will actually be used. Thus all material received by a TV station must be catalogued and stored in archives. On average only a small percentage of the archived material is actually used. The retrieval of material from such libraries is completely dependent on the quality of the cataloguing procedures.

To-day the quality of digital videos equals or even surpasses that of broadcast quality analogue material. Digital videos have the added advantage that they can be stored in media databases and be transferred without loss of quality to transmitting stations. Such digital archives can be made accessible to many TV stations, thus spreading the maintenance cost. Digital repositories furthermore open up the possibility to use content based video and image retrieval techniques. Thus individual stations need retrieve only the material they actually need for particular news casts. The net result of these advantages is that digital repositories are increasingly replacing analogue ones.

In [1] a world wide NewsGrid was proposed. The basis of the grid is a distributed digital archive, supplying versatile services for storing, searching and retrieving, downloading, and post-processing of news videos. It completely changes the way in which news video material is distributed, as well as the workflow for news editors compiling news casts. The video content retrieval process can, in addition to manual indexing techniques, employ dynamic feature extraction techniques. This reduces the cost of indexing videos manually. A disadvantage is that it requires high-speed computing resources.

Within the NewsGrid framework news videos are not distributed to subscribing TV stations. Instead news suppliers make available information about the content of available material. The users, i.e. TV stations, retrieve the desired material on a subscription or pay per item basis directly from the news suppliers' video repositories from all over the world. This results in huge savings in bandwidth and storage capacity compared to current procedures as only the material actually needed needs to be distributed. Users will, however, need fast and effective search mechanisms to retrieve specific news material.

The major potential advantages for news agencies are a huge reduction in information broadcast costs and improved statistics on which news items are actually used by clients. The use of the proposed NewsGrid requires fundamentally new marketing and distribution strategies by news suppliers. Instead of distributing news content only, information about available news material will have to be compiled and distributed or made available on the Internet.

Essential requirements for the implementation of the proposed NewsGrid are:

- Reliable and easy to use grids comprising powerful database servers
- User friendly, effective, efficient and interactive retrieval mechanisms

- Compact information being made available on news – or for that matter other – content contained in supplier databases.

2. Compilation of Newscasts

TV news casts are compiled by combining new video material received from news agencies, own correspondents or correspondents of other TV stations, with historic or reference material retrieved from archives. The classic approach is to record news videos in analogue broadcast quality on magnetic tapes. Each tape is manually catalogued with key words and content descriptions and stored in a local library. The retrieval of video material is achieved by using the catalogue and a final visual screening of selected items. This is a highly inefficient, labour and cost intensive and often ineffective process as all these actions require human actions. This holds for both the cataloguing and retrieval processes.

In the case of digital video material the repository may be realised with a DBMS (Database Management System) adapted to video content storage and retrieval, e.g. OVID [2]. The retrieval is achieved by use of catalogued descriptions and/or key words. The screening of video material for final selection can be considerably simplified [3].

Digital video databases furthermore allow the application of multimedia technologies utilising content based retrieval (CBR) methods. CBR, also referred to as feature extraction methods, have the potential of making the time and cost intensive task of manually indexing each video clip superfluous.

3. NewsGrid

The NewsGrid concept (Fig. 1) described in [1] is based on the use of dynamic feature extraction. This makes it possible to employ powerful and flexible retrieval techniques to search for particular video content. The main advantages offered by NewsGrid are that:

- News suppliers need not distribute their material to users as they can directly access and search suppliers' video repositories
- Users need only pay for material they actually use
- News agencies can gain better insight in what material is actually used by which user(s) and can thus improve their news collection activities.

Disadvantages of this approach are:

- The fact that CBR methods are still objects of research; presently available methods are promising, but still limited in their effectiveness
- The workflow of news editors is changed and they must learn to use grid technologies
- The marketing strategies of news suppliers must be adapted.

For the NewsGrid model to become accepted in the market place it is thus essential that:

- Effective CBR methods are available
- The high computational demands of CBR methods are met
- Suitable grid services to support users are available.

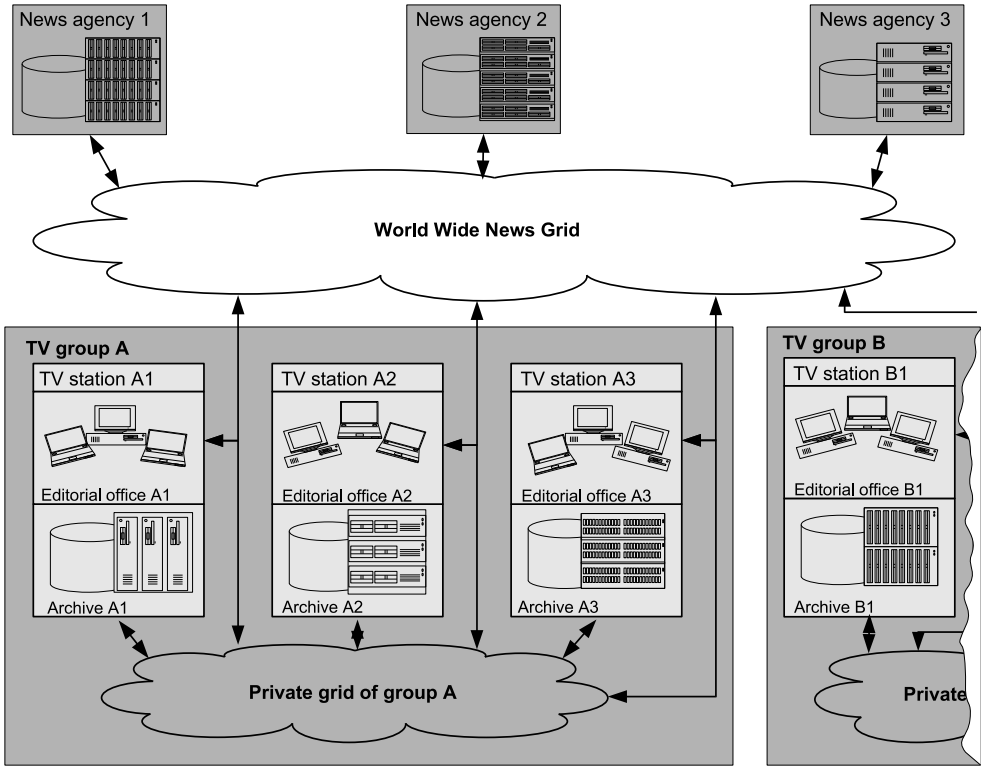

Figure 1. Structure of the NewsGrid. Several private grids (intragrids) of TV station groups are connected to the world wide NewsGrid.

To cope with the high computational demands, efficient and effective content based search algorithms are required. First results were discussed in [1]. The processing infrastructure must make available powerful parallel processing resources to achieve acceptable response times, as well as grid services that support users in their quest for information. In [1] an outline of infrastructure requirements as well as parallel processing models was discussed. In this paper the NewsGrid infrastructure and the required grid services will be explored in more detail.

4. Infrastructure

In [4] different techniques were compared in order to formulate fast parallel search methods and suitable parallel architectures. The results were obtained through simulation of the performance of various prototype configurations. The best performance is gained with a cluster of dual processor nodes and coarse-grained parallelisation. An important aspect is the communication networks connecting nodes of clusters in such a way that efficient video retrieval can be attained by a large number of users.

Within a NewsGrid news videos are stored in a relatively small number of distributed archives, thus substantially reducing redundancy. This greatly reduces the overall cost of distributing, indexing and storing archive material. The indexing process can thus be better formalised resulting in more consistent results.

The infrastructure of the whole grid is organised hierarchically (fig. 1). Each TV station has its own archive, and each TV group can have its private grid. This saves communication bandwidth. Within a smaller group higher network capacities can be implemented while the world wide grid

usually offers a lower transfer speed. Additionally downloading the same video several times should be avoided to save time and reduce costs.

Every time a participant of the grid obtains a video from another TV group or a news agency the downloaded file is temporarily stored in a local archive. If the video is accessed a second time from the same station or group, the grid software detects the video file in the local cache. Before starting a download, it is checked, whether the original material has been updated or changed. If not, the download is started from the local media server instead of from remote. The file is deleted from the cache, if it has not been accessed for a longer period. The relevance of a video evaluated by humans or automatically determined number of accesses may be used to implement a replacement policy.

A requirement for the successful implementation and use of a NewsGrid is that a set of standard methods and protocols must be agreed and adhered to. Present grid standards as well as MPEG-7 and MPEG-21 form a good starting point for defining such standards. In [5] a scheme was proposed that allows the management of news feeds stored in a cluster based database. This scheme can also be used as a basis for defining schemes that can be used in NewsGrid.

It is assumed that a news header, containing static and dynamic information, is supplied. MPEG-7 can be used to achieve this. It is also possible to supply such descriptions as separate information added to any stored videos. In general the static information supplied includes, for example Date, Time, Title, Keywords/Summary, Source, Copyright/Price, Characteristics for content based retrieval, e.g. colour moments of key frames, file name or URL of video file, etc. Additionally dynamic information can be added, such as sample sketches and images, definition of regions of interest (ROI), etc. Such descriptions can be used to define a search vector or Object Description Vector (ODV) that is defined for each video clip stored in the database. The search vector can be enhanced by adding key words and textual descriptions. These must be manually added when the video clip is stored in the database.

5. Video Storage and Retrieval

With the increasing number of available news videos the available methods using static data to define object description vectors (ODVs) are becoming increasingly inadequate for information retrieval tasks. The addition of textual descriptions and/or key words is becoming too costly. The storage and efficient and effective retrieval of video files were and are thus investigated in a number of research projects, for example VideoQ [6] and CueVideo [7]. The goal with all of these is to implement a content based search on the video data in order to avoid the expensive and inaccurate compilation of textual descriptions.

In the ODV metadata, for example title, author or date, and extracted features, such as colour, textures, shapes or motion can be stored. The presently available systems have in common that the ODVs are calculated when the video is stored in the database. As the provider cannot know what the interests of the later users will be, it is likely that significant information for the users will not be extracted and stored.

An alternate approach is to store video clips with a limited ODV containing only static data (SODV). The dynamic components of the ODV (DODV) are then calculated when the user submits his query to the video database. Thus the ODV, comprising the SODV and DODV, can only be calculated when the query of the user is known. This is an extension of a proposal by Kao for image databases [8].

Objects or persons can be found in video clips by a template matching algorithm that is performed on each image in each video clip in the database. Such systems need a huge amount of computational power. This can only be achieved with parallel systems. In the case of the proposed NewsGrid this

implies that the individual nodes accessing the video databases must have sufficient compute power available. Due to the complex and compute intensive retrieval methods this also requires that a significant parallel compute power be available across the grid.

6. The retrieval process

Instead of key words and logical operators a query can consist of a sketch or a sample image supplied by the user when he defines his search. Additionally the user can mark a region of interest (ROI), which is in most cases a figure or an object in an image. The objective is not to find exact matches, but rather to locate a set of video clips or shots that contain a frame or frames corresponding as closely as possible to the sample sketch, image or the specified region. The comparison between the query image and a video frame results in the same process as in the case of searches in image databases. In the case of an object search, template matching is used. This means the query template is compared to every equally sized region of the video frame with the same size.

7. Content based video coding

Within the NewsGrid content based video coding is used in specific situations to reduce the bitrate. This means that videos are not coded on a frame by frame basis. Instead, different objects in the video scenes are coded separately and labeled with entity identifiers.

Content based video coding can be applied to TV news casts. In the case of most TV news casts scenes are composed of a background, which never changes, an information board, which changes only at the beginning of a new theme and a speaker. The speaker is an animated object. Using this information about the scene elements a segmentation of a news cast video is possible.

In order to reduce the bitrate the background is encoded only once at the beginning of the video file, while each information board is encoded once at the beginning of each theme. Only the speaker area, which usually fills about half of the screen must be encoded with the full frame rate. Using this technique the number of pixels to be encoded can nearly be halved without any loss of quality.

A further reduction can be achieved e.g. for transmission to mobile devices. As the user is normally not interested in the background the resolution of this can be reduced by a factor of 4. Small details of the speaker are also less important, therefore the resolution can be lowered by 2 in this area. With this method the quality is only lowered in areas with low visual information and the bitrate can be reduced by over 70%. An additional aspect of mobile devices are the small displays. Content-based video coding can be used to improve readability of text by introducing user interaction. If the user selects the information board in can be displayed full-screen and the font size is increased. The content based video coding for mobile devices is described in detatil in [9].

The content based video coding can also be used for annotation and link generation. As a headline is part of the information board a text extraction can be applied to this video object. The extracted text can be used as initial key word set of the news caster scene. Similarity of information boards of different news casts can be used to automatically find candidates of related videos.

For all these use-cases additional computational power is needed. Therefore the transcoding and post-processing has to be performed on grid-ressources with high computational power.

8. Grid Services

Access to the video files made available on the repositories of NewsGrid is realised by a number of grid services [1]. In addition to the standard NewsGrid services each user must be able to start

his own programs for processing data available in repositories to which he has access. By defining different user groups and domains particular services can be made available to a user group(s).

A number of sample services needed in the case of NewsGrid are outlined here. The functionality of each service is merely exemplary and can be adapted to meet the needs of particular users:

1. File-Transfer (Download): This service enables the transcoding and downloading of a video file in combination with the specification of start and end positions.
Input parameters: Video-Id, start- and end-position, video-codec (e.g. MPEG-1/2/4, H.264), audio-codec (e.g. MP3, AAC), quality/bitrate, transmission standard (NTSC/PAL/HDTV).
Function: Recode video segment, copy recoded video to ftp-server, transmit from ftp-server to client
Result: Video file.

2. File-Transfer (Upload): With the aid of this service a new video can be uploaded and stored in the database. It is devided in the following three sub-services.

2.1. Prepare Upload: A service to prepare the upload of the new video file.
Input parameters: Author, date, time, location, title, key words, copyright.
Function: Generate database entry for new video, generate file name for new video.
Result: File name for new video.

2.2. Upload new video file: Transmit the new video file to the grid archive.
Input parameters: Video file, file name generated from sub-service prepare upload.
Function: Transfer video file via ftp from client to server, extract and classify video objects, transcode video, copy file to grid directory.
Result: Video-Id.

2.3. Generate description vector: Generates the SODV.
Input parameters: Video-Id, priority.
Function: Calculate automatically retrievable features (e.g. OCR from information board), send message to archive administrators to generate content description (using priority).
Result: Success/error.

3. Video streaming (video-on-demand service): A video is transferred to a user not with the purpose of downloading the file, but rather to view and inspect the contents.
Input parameters: Video-Id, start- and end-positions, video-codec (e.g. MPEG-1/2/4, H.264), audio-codec (e.g. MP3, AAC), quality/bitrate, transmission standard (NTSC/PAL/HDTV).
Function: A copy of the video is moved to the streaming server that manages the communication with the client. Before transmission the video is transcoded into the desired format. The user must be able to start, pause, and spool forwards and backwards, etc. as is the case with a normal video player. The appropriate parameters are made available by the video-on-demand service.
Thus only those sections of the video file to be viewed need be transferred. As the purpose is to give the user a quick insight into the video contents, the quality level can be comparatively low, which can be achieved through a stronger compression. This also reduces the amount of data to be transferred.
Result: Video stream.

4. Addition of a Description: This service is used by the archive administrators to manually add additional metadata and descriptions as well as references to related material in the database after a message from service 2.3 was received or when an update must be executed.
Input parameters: Video-Id, data for the description.
Function: Store the changed SODV in the database.
Result: Success/error.

5. Data retrieval using static features (SODV): This service enables the classic data retrieval tasks using one of the static features from the ODV. This can be directly supported by the user e.g.

in the case of key words or calculated on the client or on the server side, if the query by example technique is used. By combining different features and searching on former result sets the search can be limited to subsets of the database.
Input parameters: Feature-Id, corresponding feature vector or example video/frame.
Function: Search task is performed within the local grid and transmitted to the other nodes of the world-wide NewsGrid. Each node generates a list of the best fitting videos. These separate lists are combined to a single one.
Result: Set of results ordered by relevance.

6. Data retrieval using dynamic features (DODV) provided by information supplier: A dynamic content based search is enabled by this service. The search algorithm is made available by the information supplier, e.g. a news agency. As in the case of a combination with static data the search can be limited to a subset of the database.
Input parameters: Algorithm-Id, video object type, example video/frame, algorithm specific data, set of videos to search
Function: Search all video objects of specified type in the subset by comparing their DODV's to the example's (see section 5).
Result: Search job-Id (immediately), set of results ordered by relevance (after search is completed)

7. Data retrieval using user defined dynamic features (DODV): This service is very similar to the previous one with the difference that the algorithm to compare a video with the provided example is defined by the user. Therefore instead of the algorithm-Id an executable is used as input parameter.

8. Monitoring the search progress: Information about the progress of a search in the case of compute intensive content based searches is supplied. It must be possible for the user to ask for the display of intermediate results using the video streaming service.
Input parameters: Search job-Id.
Function: Ask every involved node to estimate the remaining time for the search and calculate overall remaining time.
Result: Overall remaining time.

9. Add new static feature for manual description: This service allows for a new static feature for manual description to be added to all videos in the database.
Input parameters: Name of the feature, data type (using MPEG-7-schema), default value.
Function: A new entry for each video in the database is created and filled with the default value. The schema for new videos is adapted. A message to the archive administrators is created to provide valid entries for the new field.
Result: Success/error.

10. Add new static feature for automatic description: Similar to the last service, but a feature for automatic description generation is added.
Input parameters: Name of the feature, data type (using MPEG-7-schema), algorithm to calculate the feature (executable).
Function: The algorithm to calculate the features is started for every video and the result will be stored in the database.
Result: Success/error

11. Video cutting: This service enables users to extract various cuts from different videos and join these to form a single new video. This is especially then the case when topical videos are to be combined with archived material. The video cutting process can thus be executed directly on the data extracted from the repository. The advantages of this service are twofold. In the first instance the amount of data to be downloaded can be substantially reduced. Secondly, news editors who only

have a moderately powerful computer, such as a notebook, available can execute cutting tasks on a more powerful grid node.
Input parameters: Video-codec, audio-codec, quality/bitrate, transmission standard and for each video to be part of the result: video-Id, start and end-position, fading-technique.
Function: Cut the video segments from the original videos and merge them to one video using fading techniques. Merge SODVs to one single SODV. Generate video-Id for the result video.
Result: Video-Id.

9. Conclusions

In this paper an infrastructure for and services supplied by a world wide NewsGrid is described. The distributed digital archive, supplying versatile services for storing, searching and retrieving, downloading, and post-processing of news videos, change the workflow for news editors. The users of NewsGrid obtain access to videos from all over the world, from any TV station or news agency connected to the grid. The data is transmitted only on demand, which saves bandwidth and storage capacity compared to current procedures.

Future work includes the development of specialised user clients, e.g. for mobile devices. The design of the services will be more detailed and more services will be added to the system. The transferability to other use cases will be improved. Furthermore concepts for digital rights managements and user roles must be defined.

References

[1] S. Geisler, G. Joubert: NewsGrid. L. Grandinetti (Ed.): Grid Computing: The New Frontier of High Performance Computing in Advances in Parallel Processing, Vol. 14, Amsterdam, etc.: Elsevier. to be published 2005 (The paper can be downloaded by the reviewers from www.in.tu-clausthal.de/~sgeisler/newsgrid/paper1.pdf using user name parco2005 and password malaga)

[2] E. Oomoto, K. Tanaka: OVID: Design and Implementation of a Video Object Database System. IEEE Trans. on Knowledge and Data Engineering, 5:629643. 1993

[3] G. Falkemeier, G. Joubert, O. Kao: Internet Supported Analysis and Presentation of MPEG Compressed Newsfeeds. Intern. Journal of Computers and Applications, Vol. 23(2), 129-136, ACTA Press. 2001

[4] S. Geisler: Efficient Parallel Search in Video Databases with Dynamic Feature Extraction, Parallel Computing: Software Technology, Algorithms, Architecture And Applications: Proc. of the Intern. Conference Parco2003, 431-438, 2004.

[5] S. Geisler, O. Kao: Cluster-Based Organisation and Retrieval of Newsfeed Archives. Proc. of the Intern. Workshop on Intelligent Multimedia Computing and Networking (IMMCN'2002), 1033-1036. 2002

[6] S.-F. Chang, W. Chen, H. Meng, H. Sundaram, D. Zhong: VideoQ: An Automated Content Based Video Search System Using Visual Cues. Proc. of ACM Multimedia, 313324. 1997

[7] D. Ponceleon, S. Srinivasan, A. Amir, D. Petkovic: Key to Effective Video Retrieval: Effective Cataloging and Browsing. Proc. of ACM Multimedia, 99107. 1998

[8] O. Kao, S. Stapel: Case study: Cairo a Distributed Image Retrieval System for Cluster Architectures. In T.K. Shih (edt.), Distributed Multimedia Databases: Techniques and Applications, 291303. Idea Group Publishing. 2001

[9] S. Geisler: Content-based Video Coding of Newscasts for Mobile Devices. International Conference on Imaging Science, Systems, and Technology (CISST 2005), 109-115. 2005